植 物 分 类 学 报 32 (4)：380—388 (1994)
Acta Phytotaxonomica Sinica

分支分类学中和谐性概念与和谐性分析方法*

徐克学

（中国科学院植物研究所系统与进化植物学开放研究实验室，北京 100093）

THE COMPATIBILITY CONCEPT AND COMPATIBILITY ANALYSIS METHOD IN CLADISTICS

Xu Ke-xue

(*Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Academia Sinica*, Beijing 100093)

Abstract Compatibility is a fundamental concept for cladistics. In this paper a mathematical definition of compatibility, called Kexue compatibility, is proposed, and then a new compatibility analysis method for cladistics based on the Kexue compatibility is developed.

At first we discuss the concept of compatibility in terms of set theory. Two sets A and B are of compatibility if A and B are disjoint, i.e. if $A \cap B = \varnothing$ (empty set), or if there exist inclusion relation, i.e. $A \subseteq B$ or $B \subseteq A$. Two classes G and H are of compatibility if any pair of sets from G and H are of compatibility, i.e. if sets A and B are of compatibility for any $A \in G$ and $B \in H$.

Let $X = [x_{ij}]$ $(i=1, 2, \cdots, t;\ j=1, 2, \cdots, n)$ is original data matrix with t OTUs $\times n$ characters. For a fixed character j we give following notations to denote its corresponding sets and corresponding class:

The corresponding sets $D(i, j) = \{k \mid x_{kj} \geqslant i, 1 \leqslant k \leqslant t\}$. $(i=0, 1, 2, \cdots, m_j)$

The corresponding class $G(A_j) = \{D(1, j), D(2, j), \cdots, D(m_j, j)\}$,

where $m_j = max\ \{x_{kj}\}$. $1 \leqslant k \leqslant t$.

Then, we introduce the definition of Kexue compatibility for cladistic character: Two cladistic characters are compatible if its corresponding classes are compatible. Let $K(i, j)$ be a characteristic function which denotes the compatibility of characters i and j, i.e.

$$k(i,j) = \begin{cases} 1 & \text{as characters } i \text{ and } j \text{ are compatible,} \\ 0 & \text{otherwise.} \end{cases}$$

---

* 国家自然科学基金资助项目（38970070）。
1992-04-28 收稿。

Let i and j run from 1 to n characters we obtain the compatibility matrix. A quantitative method for compatibility analysis is designed by means of the compatibility matrix and various numerical indexes related to the compatibility. A practical example from the evolution of animals is provided as evidence for this method.

Finally, its application to the research of cladistics is discussed.

Key words Cladistics; Compatibility

摘要 和谐性是分支分类学中的一个基本概念。本文给出一个和谐性的数学定义，称为 Kexue 和谐性。并在 Kexue 和谐性的基础上开发出一个新的和谐性分析方法。

首先在集合论的意义下讨论和谐性概念。两个集合 A 与 B 若非相交，即 $A\cap B=\varnothing$（空集），或者存在包含关系，即 $A\subseteq B$ 或 $B\subseteq A$，则称 A 与 B 是和谐的。两个集合类 G 和 H 若 G 与 H 中的任何一对集合都是和谐的，即对于任何 $A\in G$ 和 $B\in H$，A 与 B 是和谐的，则称集合类 G 与 H 是和谐的。

令 $X=[x_{ij}]_{t\times n}$ 是一个 t OTU$_s\times$n 性状的原始数据矩阵。对于固定的性状 j，我们给出下面的符号表示其对应集和对应集类：

对应集 $D(i, j)=\{k|x_{kj}\geqslant i, 1\leqslant k\leqslant t\}$ $(i=0, 1, \cdots\cdots, m_j)$

对应集类：$G(A_j)=\{D(1, j), D(2, j), \cdots, D(m_j, j)\}$

其中：$m_j=\max\{x_{kj}\}$ $1\leqslant k\leqslant t$

于是，我们引导出分支性状的 Kexue 和谐性定义：两分支性状的对应集合类是和谐的，则称该分支性状相互是和谐的。取特征函数 K (i, j) 表示性状 i 与 j 的和谐性，即：

$$K(i,j)=\begin{cases}1 & \text{当性状 } i \text{ 与 } j \text{ 和谐，}\\ 0 & \text{否则。}\end{cases}$$

令 i 与 j 跑遍从 1 到 n 性状，得到和谐性矩阵。借助和谐性矩阵和各种与和谐性有关的数量指标，设计和谐性定量分析方法。来自动物演化的一个实际例子为此方法提供验证。

最后，对该方法在分支分类研究中的应用进行讨论。

关键词 分支分类学；和谐性

定量分支分类研究，从性状的演化关系导出生物的演化过程，要求性状之间保持和谐性。因而性状之间的和谐性是分支分类学的重要概念。Camin 和 Sokal (1965) 首次提出和谐性概念，从这个概念引出分支性状和谐性的计算方法。可是 Camin 与 Sokal 给出反映和谐性的计算值竟然是不对称的。即性状 p 对性状 q 的和谐性值 $C(p, q)$ 与性状 q 对性状 p 的和谐性值 $C(q, p)$ 不相等，这显然不符合通常和谐性的正确理解。当时由于未能对和谐性给出严谨且合理的定义，自然不可能得出正确的计算方法。Estabrook 等 (1976) 给出和谐性的一个数学定义，这个定义克服了不对称性的缺陷。该定义建立在分支性状的“真实性”意义上，所谓“真实”是从抽象代数中提出假设的三项条件：

(1) $\bar{a}\in K^{-1}(K(a))$，在此 $\bar{a}=\wedge K^{-1}(K(a))$；

(2) 若 $a\leqslant b$ 则 $K(a)\leqslant K(b)$；

(3) 若 $K(a)\leqslant K(b)$ 则 $\bar{a}\leqslant\bar{b}$.

这里 K 是分支单位到分支性状态的映象，a 与 b 是任意两个分支单位。

Estabrook 的三项“真实性”条件要求十分苛刻，内容抽象，不易理解，更难于联系生物演化的实际问题。Estabrook 的理论在分支分类学中颇有影响，但是他的和谐性定义

至今未被广泛采用。为了得到一个合理的、有实际意义的定义，并探寻对性状进行和谐性分析的方法，不得不另辟蹊径。

（一）从集合的和谐性到生物演化性状的和谐性

在此，先从集合论引入两个集合和谐性概念。我们给出集合之间和谐性的定义：

两集合 A 与 B，若满足以下两条件之一： 条件 1. $A\cap B=\varnothing$（空集）； 条件 2. 若 $A\cap B\neq\varnothing$，则 $A\subseteq B$ 或 $B\subseteq A$，则称集合 A 与 B 和谐。

例如四个集合 $A=\{1, 2, 3, 4, 5, 6, 7, 8\}$；$B=\{9, 10\}$；$C=\{7, 8, 9, 10\}$；$D=\{1, 2, 3\}$。其中 B 与 D，C 与 D 和谐，因为它们的交集为空集，集合 A 与 D，B 与 C 和谐，因为它们之间存在包含关系，$D\subseteq A$，$B\subseteq C$； 集合 C 与 A 不和谐，因为 $C\cap A=\{7, 8\}$ 非空，且彼此之间不存在包含关系。

以集合为元素，许多集合构成集合类，我们把集合之间的和谐性引到集合类上去。集合 X 的部分子集合构成集合类，如此两个集合类 G 与 H，若分别从 G 和 H 中取出的任意两个集合都是和谐的，则称集合类 G 与 H 是和谐的。与自身和谐的集合类称为和谐的集合类。

例如三个集合类 $G=[\{a\}, \{b\}, \{c\}, \{b, c\}, \{a, b, c\}]$，$H=[\{a\}, \{b\}, \{c\}, \{a, b, c\}]$ 和 $F=[\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, b, c\}]$。容易验证分别从 G 和 H 中提取的两集合都是和谐的。因而集合类 G 与 H 和谐。与此相同，集合类 H 与 F 和谐。但是 G 与 F 不和谐，因为 G 中存在集合 $\{b, c\}$，F 中存在集合 $\{a, b\}$，而 $\{b, c\}$ 与 $\{a, b\}$ 两集合不和谐。

再把前述和谐性概念引伸到分支性状的和谐性。为此假设一个分支分类原始数据矩阵，表示如下：

$$t\text{ 个分类单位 OTU}\quad \overset{n\text{ 个性状}}{\begin{bmatrix} x_{11} & x_{12}\cdots & x_{1n} \\ x_{21} & x_{22}\cdots & x_{2n} \\ \cdots & \cdots & \cdots \\ x_{t1} & x_{t2}\cdots & x_{tn} \end{bmatrix}}$$

这里要求每个性状状态的编码数值均取非负整数，数值自 0 开始，从小到大数值的增加与演化方向一致。

如果 x_{ij}（$i=1, 2\cdots, t$；$j=1, 2, \cdots, n$）表示第 i OTU 第 j 性状的编码数据；m_j 表示性状 j 最大状态编码值。把分类单位中第 j 性状编码值大于等于 i 的分类单位集合记作 $D(i, j)$，该集合表示如下：

$$D(i, j)=\{k|x_{kj}\geqslant i, 1\leqslant k\leqslant \mathrm{t}\}\qquad (j=1, 2, \cdots, n;\ i=0, 1, 2, \cdots, m_j)$$

这里 k 代表分类单位的编码。我们可以把数值 k 看作就是第 k 个 OTU，这样集合 $D(i, j)$ 既是分类单位的编码集合，也是相应的分类单位集合。在演化集合的理论中有最近共同祖先的概念，两个分支单位 x 与 y 的最近共同祖先记作 $x\wedge y$，分支单位集合 S 的最近共同祖先以 $\wedge S$ 表示。如果 $D(i, j)$ 的最近共同祖先以 a_{ij} 表示，可写成：$a_{ij}=\wedge D(i, j)$. 在生物演化关系中，如果 $a_{i,j}$ 是分支单位的祖先，表示成 $a_{ij}\leqslant y$，那么，对于任意分

支单位 a，我们把分类单位集合的一个子集

$$C(a) = \{y \mid a \leqslant y, y \in O\ (\text{分类单位集合})\}$$

称为 a 的演化对应集。演化集合论的理论可以证明，任何集合 $D(i, j)$ ($1 \leqslant j \leqslant n$, $0 \leqslant i \leqslant m_j$)，其最近共同祖先的演化对应集就是集合 $D(i, j)$ 自身，即： $D(i, j) = C(a_{ij})$。因为这个缘故我们可以把 $D(i, j)$ 也称为演化对应集。

对于一个固定的性状 j ($1 \leqslant j \leqslant t$)，让 i 跑遍除 0 以外所有可能的状态值 ($1 \leqslant i \leqslant m_j$)，得对应集 $D(1, j)$，$D(2, j)$，…，$D(m_j, j)$，以及相应的最近共同祖先 a_{1j}，a_{2j}，…，$a_{m_j j}$。我们把这些最近共同祖先记作

$$A_j = \{a_{1j}, a_{2j}, \cdots, a_{m_j j}\} \qquad (j = 1, 2, \cdots, t)。$$

相应的对应集合类记作 $G(A_j) = [D(1, j), D(2, j), \cdots, D(m_j, j)]$ ($j = 1, 2, \cdots, t$)。

在集合类 $G(A_j)$ 中，任意两个集合 $D(i_1, j)$ 和 $D(i_2, j)$，如果 $0 \leqslant i_1 \leqslant i_2 \leqslant m_j$，从集合 $D(i, j)$ 的定义故有

$$D(i_2, j) \subseteq D(i_1, j)$$

再从集合类的和谐性定义知，任意性状 j 的对应集类 $G(A_j)$ 自身是和谐的。然而任意两个性状 k 与 l，其相应的对应集类 $G(A_k)$ 与 $G(A_l)$ 是否彼此和谐呢？由此总结出性状和谐性的定义：

两分支分类性状，如果其演化对应集类彼此和谐，则称两性状和谐，否则称为不和谐。我们把按此意义下定义的和谐，称为 KEXUE 和谐。

现在以桔梗科的试验数据来具体说明 KEXUE 和谐性概念。6 个 OTU×8 个性状的数据 ($t=6$，$n=8$) 见表 1。

表 1 桔梗科试验数据

Table 1 A data for experiment from the family of Campanulaceae

分类单位 OTUs	性状 Characters							
	①茎缠绕否	②株高	③叶序	④叶缘	⑤花序	⑥子房室数	⑦果裂方式	⑧种具翼否
1 党参 *Codonopsis pilosula*	1	1	1	0	0	1	2	0
2 桔梗 *Platycodon grandiflorus*	0	0	0	1	0	2	1	0
3 轮叶沙参 *Adenophora pereskiifolia*	0	0	2	1	2	0	0	0
4 荠苨 *Adenophora remotiflora*	0	0	0	2	1	0	0	0
5 羊乳 *Codonopsis lanceolata*	1	1	1	0	0	1	2	1
6 石沙参 *Adenophora polyantha*	0	0	0	1	2	0	0	0

例如性状 3 的数据是 1 0 2 0 1 0 ($m_3=2$)，有两个演化对应集：

$$D(1, 3) = \{1, 3, 5\}$$

$$D(2, 3) = \{3\}。$$

性状 3 的演化对应集类是

$$G(A_3) = [\{1, 3, 5\}, \{3\}]。$$

类似地，性状 4 的演化对应集类是

$$G(A_4) = [\{2, 3, 4, 6\} \{4\}]。$$

在 $G(A_3)$ 与 $G(A_4)$ 之间有两个分类单位集合 {1，3，5} 与 {2，3，4，6} 相互不和谐，因此 $G(A_3)$ 与 $G(A_4)$ 不和谐，由此可知性状 3 与 4 不和谐。而性状 5 与性状 4 却是和谐的，因为 $G(A_5)$ = 〔{3，4，6}，{3，6}〕与 $G(A_4)$ 和谐。

（二）分支性状的 KEXUE 和谐性分析

在分支性状之间建立新的和谐性定义之后，就可以对性状进行和谐性分析。先建立函数 $K(i, j)$：

$$K(i,j)=\begin{cases}0 & \text{当性状 } i \text{ 与性状 } j \text{ KEXUE 和谐}\\ 1 & \text{当性状 } i \text{ 与性状 } j \text{ 非 KEXUE 和谐}\end{cases}$$

然后再确立描述性状 i 不和谐程度的性状不和谐数 $K(i)$：

$$K(i)=\sum_{j=1}^{n} K(i,j)$$

以及性状 i 不和谐系数：

$$C(i)=\frac{1}{n-1}K(i)。$$

对整个数据和谐性的评估，有不和谐总数：

$$K=\sum_{i=1}^{n} K(i),$$

以及不和谐总系数：

$$C=\frac{K}{n(n-1)}$$

在 KEXUE 和谐性分析中，性状不和谐数 $K(i)$ 与不和谐总数 K 是性状之间不和谐联系的计数。0 值表示完全和谐，值愈大表示和谐性愈差。性状不和谐系数 $C(i)$ 与不和谐总系数 C 是性状之间不和谐联系占全部性状联系的比率。系数值介于 0 与 1 之间，0 值表示完全和谐，值愈大表示和谐性愈差，值 1 表示完全不和谐。在分析性状之间的和谐性时，列出全部 $K(ij)$ $(i, j=1, 2, \cdots, n)$ 值对了解性状之间的和谐性有价值。因而设计 KEXUE 和谐性分析应包括函数 $K(i, j)$ $(i, j=1, 2, \cdots, n)$ 构成的和谐性矩阵。此外还应包括性状不和谐数 $K(i)$ $(i=1, 2, \cdots, n)$，不和谐总数 K，性状不和谐系数 $C(i)$ $(i=1, 2, \cdots, n)$ 以及不和谐总系数 C。

以桔梗试验数据为例，具体说明 KEXUE 和谐性分析的计算步骤。

原始数据要求非负整数，每个性状数值从小到大体现性状状态的进化次序，0 值为最原始状态。

（1）首先按照 $D(i, j)$ 的定义列出演化对应集类 $G(A_j)$ $(j=1, 2, \cdots, n)$。（表 2 示桔梗科试验数据的对应集与对应集类）

（2）然后对任意两性状 i，j 之间的对应集类 $G(A_i)$ 与 $G(A_j)$ 进行和谐性检验，获值 $K(i, j)$ $(i, j=1, 2, \cdots, n)$。将值 $K(i, j)$ 排成和谐性矩阵。

（3）在和谐性矩阵上按行累计求和，计算每个性状不和谐数 $K(i)$ 与性状不和谐系数 $C(i)$。分析每个性状与整个数据其它性状的和谐性关系。

（4）最后计算不和谐总数 K 与不和谐总系数 C，对整个数据的和谐性进行评价。

表2　桔梗科试验数据的对应集和对应集类

Table 2　The corresponding set and corresponding class of data from Campanulaceae

性　状 Characters	对应集类 Corresponding classes	对应集 Corresponding sets
1	$G(A_1)$	{1, 5}
2	$G(A_2)$	{1, 5}
3	$G(A_3)$	{1, 3, 5}, {3}
4	$G(A_4)$	{2, 3, 4, 6}, {4}
5	$G(A_5)$	{3, 4, 6}, {3, 6}
6	$G(A_6)$	{1, 2, 5}, {2}
7	$G(A_7)$	{1, 2, 5}, {1, 5}
8	$G(A_8)$	{5}

表3　KEXUE 和谐性分析（性状比较）

Table 3　KEXUE compatibility analysis (comparison with character)

$K(i, j)$ i \ j	1	2	3	4	5	6	7	8	性状 i 不和谐数 $K(i)$ Incompatible number of charater i	性状 i 不和谐系数 C (i) Incompatible coefficient of character i
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	1	1	1	1	0	4	0.571
4	0	0	1	0	0	1	1	0	3	0.429
5	0	0	1	0	0	0	0	0	1	0.143
6	0	0	1	1	0	0	0	0	2	0.286
7	0	0	1	1	0	0	0	0	2	0.286
8	0	0	0	0	0	0	0	0	0	0

性状不和谐总数＝12，性状不和谐总系数＝0.2143。

Incompatible total number of character＝12。Incompatible total coefficient of character＝0.2143。

全部分析计算结果可总结为表3的形式。表3中列出了每一对性状之间的和谐性状况，0值和谐，1值不和谐。又从性状不和谐数 $K(i)$ 与不和谐系数 $C(i)$ 的比较中，看出性状4与性状3具有较高的值，说明这两个性状与其它性状和谐性较差。有待重新考虑性状的演化关系。至于性状不和谐总数与性状不和谐总系数是对整个数据和谐性程度的衡量值。该值亦分别类似于性状不和谐数与性状不和谐系数，具有相同的意义。

（三）性状 KEXUE 和谐性分析应用的评价

KEXUE 性状和谐性分析方法基于对性状 KEXUE 和谐性定义，演化集合论严谨地论证了该定义与生物演化的本质联系。因此构筑于其上的和谐性分析方法，能够正确地反映性状之间的和谐性关系。

按照 Camin 与 Sokal 意义下的和谐性，对桔梗科试验数据，可以得到与表2形式上类似的分析结果（表4）。

从表4可以看出，在行的方向上，性状3、6、7与其它性状具有较明显的不和谐性，而从列的方向上却显示出性状8与其它性状具有不和谐性。由于 Camin-Sokal 意义下的和谐性数值表示不对称，导致行与列方向上分析结果的不一致。因此基于 Camin-Sokal 和谐性意义的和谐性分析不能得到正确的分析结果。

表 4　Camin-Sokal 和谐性分析

Table 4　A result of compatibility analysis by Camin-Sokal method

C (h, i) ＼ i ／ h	1	2	3	4	5	6	7	8	和谐值 Comapatibilities	附加步数 Extra steps
1	×	0	0	0	0	0	0	1	6	1
2	0	×	0	0	0	0	0	1	6	1
3	1	1	×	1	1	1	1	1	0	7
4	0	0	1	×	1	1	1	0	3	4
5	0	0	2	1	×	0	0	0	5	3
6	1	1	1	1	0	×	0	1	2	5
7	1	1	1	1	0	0	×	2	2	6
8	0	0	0	0	0	0	0	×	7	0
和谐值 Compatibilities	4	4	3	3	5	5	5	2	和谐值总数 31 Total compatibilities	
附加步数 Extra steps	3	3	5	4	2	2	2	6		附加总步数 27 Total extra steps

表 5　动物演化数据

Table 5　A date from the evolution of animals

性　　状	OTU' s														
	1 人	2 猿	3 狗	4 牛	5 马	6 鲸	7 鸟	8 蛇	9 蛙	10 鲤	11 虾	12 蝗虫	13 蚌	14 螺	15 水螅
1 体对称形式	2	2	2	2	2	2	2	2	2	2	2	2	2	0	1
2 生活方式	2	2	2	2	2	0	3	2	1	0	0	3	0	0	0
3 生殖方式	3	3	3	3	3	3	2	2	1	1	1	2	1	1	0
4 脊椎骨动物	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
5 体分节	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
6 循环方式	4	4	4	4	4	4	4	3	3	2	1	1	1	1	0
7 体恒温	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
8 呼吸方式 1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	0
9 呼吸方式 2	1	1	1	1	1	1	1	1	1	1	1	2	1	1	0
10 消化器官 1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	0
11 消化器官 2	1	1	1	1	1	1	1	1	1	1	1	2	1	1	0
12 消化器官 3	4	4	4	5	4	5	3	3	2	2	1	1	1	1	0
13 消化器官 4	3	3	3	3	3	3	4	3	2	2	1	1	1	1	0
14 神经器官 1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	0
15 神经器官 2	3	3	3	3	3	3	4	3	2	2	1	1	1	1	0
16 神经器官 3	5	4	4	4	4	4	3	3	2	2	1	1	1	1	0
17 体表类型 1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
18 体表类型 2	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
19 体表类型 3	3	3	3	3	3	3	4	3	2	1	0	0	0	0	0
20 体表类型 4	4	4	4	4	4	5	3	3	2	1	0	0	0	0	0
21 体表类型 5	5	4	4	4	4	4	3	3	2	1	0	0	0	0	0

KEXUE 性状和谐性定义中，由于函数值 $K(i, j)$ 是对称的，即

$$K(i, j) = K(j, i)$$

因而基于该值的和谐性分析计算，能正确反映出性状之间的和谐性关系。

为了进一步验证 KEXUE 和谐性分析方法在实际应用的可靠性，以动物演化的真实例子构成数据（表 5），进行和谐性分析计算。表 4 选取了从腔肠动物到高等哺乳动物计 15 个 OTU 的数据；采取了包括各主要器官，反映动物进化过程的主要性状，计 21 个。性状数据取非负整数，按演化的次序从 0 逐步增加，随着数值的增加代表性状状况进化的过程。由于这些动物的进化过程都已清楚，性状演化的极性关系也比较清楚。真正反映动物自然进化过程的数据，其性状之间应该完全和谐。为了达到对方法可靠性进行检验的目的，故意在前三个性状中安排了性状进化中的错误数据。KEXUE 和谐性分析结果列于表 6 中。分析结果十分正确地显出，前三个性状不和谐数和不和谐系数都十分突出地高于其它性状。

表 6　动物演化数据 KEXUE 和谐性分析（性状比较）

Table 6　A KEXUE compatibility analysis by comparison of corresponding set on a data of animal evolution

性　状 i	不　和　谐　记　录	不和谐数 $K(i)$	不和谐系数 $C(i)$
1	0 0 1 0 0 1 0 1 1 1 1 1 1 1 1 1 1 0 0 0 0	12	0.6000
2	0 0 1 1 1 1 1 1 0 0 0 1 1 1 1 1 0 1 1 1 1	15	0.7500
3	1 1 0 1 1 1 0 1 0 0 0 1 1 1 1 1 0 1 1 1 1	15	0.7500
4	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000
5	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000
6	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
7	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	0.0500
8	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
9	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	0.0500
10	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	0.0500
11	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	0.0500
12	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
13	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
14	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
15	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
16	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3	0.1500
17	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	0.0500
18	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000
19	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000
20	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000
21	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2	0.1000

Lmax=463 steps, Lmin=59 steps, $K=80$, $C=0.1905$

在表 5，动物演化数据前三个性状中，安排的错误是动物系统进化关系上一些有趣的例子，值得进行讨论。

第 1 个性状，体对称形式，具有不对称，辐射对称和左右对称三种状态。从不对称到辐射对称，从辐射对称到左右对称这是动物进化的趋势，三个状态的编码值分别为 0，1 和 2。而螺的体形不对称，作为复足类不对称体形起源的研究，扭转论（theory of torsin）已对此作出正确的分析。古生物学，胚胎学都证实早期复足类动物体形是左右对称的。适应于行走方面的需要，体形从左右对称到不对称是复足类动物的进化表现。数据中将螺的体形不对称，形式上取 0 值是一个错误。只有当动物学家在研究动物起源时，

正确地阐明复足类动物不对称体形的来源之后，这个错误才得到澄清。而 KEXUE 和谐性分析居然能够十分敏锐地从性状的和谐性显示出该性状存在的问题。

第2个性状，生活方式，具有水生、两栖、陆生和能飞行三种状态。从水生到陆生是脊椎动物的进化表现。然而作为哺乳动物的鲸，它的进化是从陆地再次返回水中，对该性状的演化过程，如果把鲸与鱼类同属于一个演化阶段是一个错误。又鸟类的会飞行和节肢动物蝗虫的能飞行，是在各自不同的状态下，朝着形式上相同而实质上不同的方向进化。因而把鸟与蝗虫取相同的状态编码值，在演化关系上视作完全等同，这是第2个错误。由于该性状具有两方面的错误，它比性状1具有更高的不和谐数与不和谐系数。

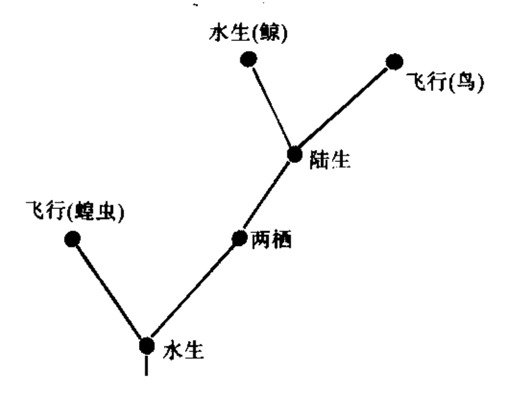


图1 动物生活方式的进化

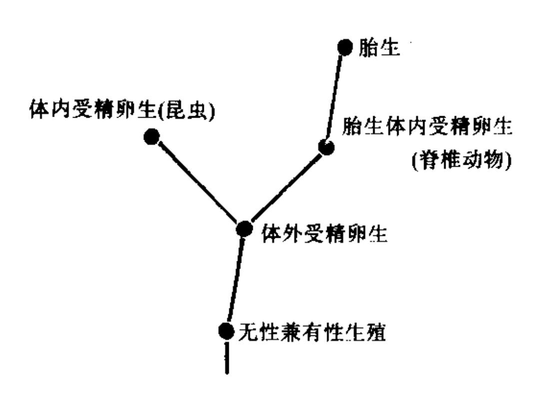


图2 动物生殖方式的进化

第3个性状生殖方式，它的演化关系如图3所示。如数据中所处理，把蝗虫的体内受精卵生和脊椎动物蛇与鸟的体内受精卵生视为完全相同的进化阶段这也是一个错误。此外本性状在虾、鲤、蛙、蛇和鸟的演化进程上也有一些不合适之处，因而造成该性状与其它性状的高度不和谐性。

定量分支分类是研究生物系统学的重要方法，受到普遍重视。然而当今生物分类学家都只追求较好的分支分类方法，而忽视对性状和谐性的定量分析。殊不知如果不能得到正确反映性状演化关系的数据，再好的定量分支分类方法也不可能得到符合真实情况的生物演化过程。本文提出的和谐性分析方法，可为分支分类研究在进行分支分类运算之前，帮助生物学家分析性状之间的和谐性，从修正原始数据的错误中获得更正确的分支分类结果。

当被分析的数据较大时，以手工进行该方法的运算已不可能，需要依靠计算机。编写如此和谐性分析方法的计算机程序并不十分困难，作者已用 BASIC 语言写出程序，在 IBM 微机上实现了 KEXUE 和谐性定量分析运算。该程序已纳入 KEXUE 定量分析软件系统，欢迎读者使用。考虑到 PASCAL 语言的集合类型可以实现对集合的直接运算，因此以 PASCAL 语言编写程序可望更容易实现，并获得更快的计算效果。

参 考 文 献

徐克学. 1989. 分支分类的一种计算方法——最大同步法. 植物分类学报, **27** (3): 232—239.

徐克学, 1993. 数量分类学 (待出版中). 北京: 科学出版社.

徐克学. 1992. 生物演化的数学模型. 生物数学学报, **7** (3): 92—97.

Camin J H, Sokal R R. 1965. A method for deducing branching sequences in phylogeny. Evolution, **19** (3): 311—326.

Estabrook G F, Johnson C S, McMorris Jr et al., 1976. A mathematical foundation for the analysis of cladistic character Compatibility. Mathematical Bioscience, **29**: 181—187.